Study of the Pathomechanism and Progression of COVID-19: Epidemiological, Diagnostic, Clinical and Intensive Care Aspects

Doctoral (PhD) Thesis

Ágoston László Hamar MD

Clinical Medical Sciences Doctoral School

Head of the Doctoral School: Prof. Lajos Bogár MD, PhD, DSc

Head of the Program: Prof. Attila Miseta MD, PhD, DSc

Supervisor:

Katalin Gombos MD, PhD, associate professor,

Department of Laboratory Medicine

University of Pécs

Medical School

Pécs, 2024

I. Introduction

SARS-CoV-2 was first detected in December 2019 in Wuhan, China. The first infected patients were hospitalized with a diagnosis of pneumonia of unknown origin, with the common link being the livestock and fish market in Wuhan. In the timespan of more than 4 years since the emergence of the virus, mortality has shown considerable fluctuation both by variant and by country. According to the WHO definition of VOC (Variant of Concern), the most prevalent in the European region were alpha (B.1.1.7), delta (B.1.617.2), omicron (B.1.1.529) and sublineages of these variants. At the time of writing, more than 704 million infections and more than 7 million deaths have been confirmed worldwide, according to worldometers.info.

The genome of SARS-CoV-2 is nearly 30 kilobases in size, encoding 29 proteins with 4 structural proteins: S (spike), M (membrane), E (envelope), N (nucleocapsid) proteins. Non-structural proteins (e.g. RNA-dependent RNA polymerase, RdRp) are involved in RNA replication, immune evasion, and additional proteins are involved in infection, survival, host-cell communication.

The gold standard method for the diagnosis of SARS-CoV-2 infection is real-time reverse transcription quantitative polymerase chain reaction (RT-qPCR), which has an exceptionally high sensitivity and specificity.

Antigen tests - most of which are based on the lateral flow immunoassay (LFIA) technique - also play an important role in diagnostics, with positivity being a reliable indicator of current infection (high specificity), while a negative test does not exclude infection (lower sensitivity). The majority of SARS-CoV-2 rapid antigen tests (RATs) are aimed at detecting SARS-CoV-2 N protein in respiratory tract specimens (including the Abbott Panbio RAT examined in this thesis).

Antibody/serological tests represent the third main group of SARS-CoV-2 diagnostics, measuring antinucleocapsid IgM and IgG during infection and anti-spikeRBD-protein IgG humoral immune response to immunization by serological methods.

The clinical spectrum of COVID-19 disease due to infection with the SARS-CoV-2 virus ranges from asymptomatic to critically severe conditions (respiratory failure, septic shock and/or multi-organ failure). Most comorbidities associated with

COVID-19 are associated with increased hospitalization and mortality risk. Among the laboratory and clinical parameters in the intensive care unit (ICU), low lymphocyte count, high total white blood cell count, elevated procalcitonin (PCT), C-reactive protein (CRP) and ferritin levels, lower PaO_2/FiO_2 ratio (Horovitz index) and higher chest CT severity score (CTSS) were the most important factors which influenced the adverse outcome.

Machine learning (ML) is is a rapidly growing field of science that is useful and capable of analyzing large amounts of complex health data and supporting medical decision making. In predicting the outcome of infectious diseases, ML models can reveal (even hidden) relationships in data that help predict disease severity, outcome, and probability of survival. Since the outbreak of COVID-19, several studies have investigated the use of ML methods to predict patients' mortality risk. ML methods can be divided into two broad categories based on complexity: shallow and deep learning methods. The former has a simpler structure, is easier to learn and contains fewer layers, which is why a large proportion of research uses these methods. Deep learning (e.g. neural networks) requires complex structures and extremely large amounts of data and computational resources, which may make it less accessible to researchers and separate health care institutions. In the literature on SARS-CoV-2 ICU research, machine learning methods belonging to the "standalone" group of shallow learning methods have been the most commonly used (Viderman et al., 2023). They are characterized by not being part of a larger algorithm, but operating independently of other models and are able to make predictions/classifications independently based on inputs. Examples of such methods are random forest (RF), logistic regression, decision tree, and support vector machine. RF is one of the most widely used ML algorithms, its popularity is due to its less tendency to "over-learn" (it also provides reliable performance in predicting new, previously unknown data, not only on the learning dataset), and also to consider many parameters simultaneously, and the reliance on a combination of multiple decision trees (Breiman et al., 2001). RF is also a major focus of research on the SARS-CoV-2 ICU patient population.

II. Major aims of the thesis

II/1. Aims regarding the dynamic, interactive epidemiological map

Researchers at the prestigious Johns Hopkins University in Baltimore, USA were the first to create a dynamic map that could track outbreaks week by week; first in different US states, later worldwide. The aim of our workgroup was to create a similar digital, dynamic, timeline-like map for the Southern Transdanubia region (Baranya, Somogy, Tolna counties), which provides more detailed information and a population-based epidemiological overview using color coding. Here we present raw and population density stratified data for tests, with a special focus on positive cases, as well as other data: average age, presence of symptoms, cycle threshold (Ct) of PCR tests, prevalence of viral variants. In this way, regional data and trends can be visualized and the basis of the map can be used for other purposes, e.g. in case of the emergence of other infectious diseases or even for regional visualization of other non-communicable diseases, if sufficient and appropriate data are available. We also aimed to analyze the identified virus variants.

II/2. Aims regarding the analytic performance of Panbio rapid antigen test

The gold standard for identifying SARS-CoV-2 virus is the RT-qPCR method, but various rapid antigen tests have also become increasingly popular. The aim of our retrospective study was to evaluate the diagnostic performance of Panbio (Abbott Rapid Diagnostics, Jena, Germany) rapid antigen test (which was the only available RAT at the University of Pécs Clinical Centre in the timeframe of our study), mainly in comparison with RT-qPCR in clinical settings.

II/3. Aims regarding the study of clinical, laboratory and molecular genetic parameters associated with adverse outcomes of SARS-CoV-2 infection

The greatest burden of the COVID-19 outbreak concentrated in ICUs. The aim of our study was to compare the three main viral variants (Alpha, Delta, Omicron) based on viral genome sequencing results, using a number of different clinical and laboratory parameters. The focus was on mortality and comorbidities. In addition, a machine learning model, the random forest (RF) algorithm, was used to predict mortality and to identify the clinical and laboratory parameters most relevant for prediction.

III. Materials and methods

III/1. Materials and methods regarding the dynamic, interactive epidemiological map

A nasopharyngeal swab was taken from patients by trained staff of the National Ambulance Service (OMSZ) or by volunteer medical/health science students, or by nursing and medical staff in hospitals. Samples were taken from the opposite side of the nasal septum, by applying slight pressure to the surface of the nasal mucosa, and then the sampling stick was washed into a tube containing virus transport medium (VTM). The sampling swab containing nasopharyngeal and/or oropharyngeal swab samples was placed in a VTM tube, tightly capped and transported in a double sealed airwrapped package at 4°C to the laboratory where the samples were processed.

The data collection covers the timespan between 2020.08.19 and 2022.02.13 and includes 271,849 COVID-19 test data from Baranya, Somogy and Tolna counties.

Data was partly recorded manually by the administrators of the Department of Laboratory Medicine on the basis of the test questionnaires and partly extracted from the local hospital IT system (e-MedSolution, T-Systems, Hungary). The final database was stored in an anonymized Excel spreadsheet (Microsoft, Redmond, WA, USA) with the following data: date of testing, date of birth, age, sex, postcode, county, place of residence (municipality), PCR test result (positive/negative), presence of symptoms (yes/no), cycle threshold value of positive PCR test, and type of viral variant if viral genome sequencing was performed. The anonymous data was processed and visualized on a website in a joint project with PCSUNIQ Kft., final results are available at https://covid-pte.vercel.app/dashboard. The incidence of positive cases (infection) per municipality is color-coded: all positive cases in a municipality are divided by the population per 100,000 inhabitants. Municipalities are displayed on a color scale from green to deep red based on 0-20%, 20-40%, 40-60%, 60-80% and 80-100% of the maximum value. At the bottom of the webpage, you can view and search the total number of tests, positive cases, complaints and no complaints per municipality. The main viral variants (European, Alpha, Delta, Omicron) are also available by municipality and case number at https://covid-19-spread-map-nuxt.vercel.app/.

III/2. Materials and methods regarding the analytic performance of Panbio rapid antigen test

Our retrospective study analyzed data gathered in the time spanning from 21 January 2021 through 30 April 2021. A total number of 5,136 parallel Panbio RADT and RT-qPCR samples were included from all departments of the Clinical Center, University of Pécs, Hungary. During the study protocol, all patients tested with Panbio RADT were tested in parallel by for SARS-CoV-2 RT-qPCR. Inclusion criteria were the presence of a SARS-CoV-2 RAD Panbio test result combined with an RT-qPCR test result, both performed within 24 hours. Panbio RAD tests were performed and evaluated by trained health care professionals. The diagnostic PCR tests were carried out in the Department of Laboratory Medicine in full accordance to a protocol accredited by the National Accreditation Authority (NAH-9/0008/2021, L7/6 MLMB 06 2020.4-1).

Due to the limited space available in this thesis, a detailed description of the sample collection and RT-qPCR assay is given in the full dissertation.

A total number of 5,136 samples tested parallel for SARS-CoV-2 Panbio RADT and SARS-CoV-2 RT-qPCR were collected from 4,440 individuals who were admitted in the clinical departments of the University of Pécs, Clinical Center. The test results and demographic data were originally documented in the local hospital information system (e-MedSolution, T-Systems, Hungary). Our extracted data was registered using Excel 2015 (Microsoft, Redmond, WA, USA). The final database includes an anonymized ID from both name and insurance number. It also contains information referencing gender, age, time and place (department) of test, RAD and PCR test result, Ct value, presence of symptoms and number of days until a negative PCR test in the event of a previous positive PCR test and mortality.

III/3. Materials and methods regarding the study of clinical, laboratory and molecular genetic parameters associated with adverse outcomes of SARS-CoV-2 infection

Our retrospective study consists of a collection of 503 clinical isolates of SARS-CoV-2 with clinical parameters and laboratory biomarkers plus diagnostic values collected from 503 ICU patients in a timeframe between January 2021 through November 2022 when Alpha (B.1.1.7), Delta (B.1.617.2), and Omicron (B.1.1.529)

and their sub-lineages were circulating SARS-CoV-2 lineages in the Southern Transdanubian region, Hungary. ICU patients were included who had positive SARS-CoV-2 RT-qPCR test results and were admitted due to COVID-19. Exclusion criteria were SARS-CoV-2 positive patients who were admitted to the ICU mainly due to other reasons than complications associated with COVID-19 (polytrauma, traumatic brain injury, diabetic ketoacidosis, etc.), patients with hematologic malignancies and other SARS-CoV-2 isolates such as B.1.160 (20A.EU2). 201 clinical isolates produced results of viral WGS, and the remaining 302 clinical isolates were included by generating a 95% range estimation based on the regional sequencing database, which was made for the dynamic, interactive epidemiological map mentioned in the first part of my thesis. Originally, we began with 510 ICU patients, however, seven patients were excluded in the transition phase of Delta to Omicron, since the range estimation was not possible due to the overlap between the two variants in the last week of December 2022 and the first week of January 2023.

Patient clinical and demographic data were originally documented in the local hospital information systems (e-MedSolution {T-Systems, Hungary}, IntelliSpace Critical Care and Anesthesia [ICCA] {Philips Medical Systems, USA}). The viral genomic results were originally stored in the Genomics and Bioinformatics Core Facility, Szentágothai Research Centre. Our extracted data was registered using Excel 2015 (Microsoft, Redmond WA, USA). The final manual database includes an anonymized ID from name and insurance number. The database also contained information regarding the following: gender, age, date of testing, RT-qPCR test results with the cycle threshold values, viral WGS results, days spent in ICU, chest CTSS which is calculated based on the percentage of lung lobe involvement [1–5 severity score added up per lung lobe, maximum score is 25], mortality, vaccination history, P/F ratio with two columns, one with the first value following ICU admission, and one with the lowest value during the first 24 h of ICU stay, follow-up of the ICU survivors for 28 days following ICU discharge, laboratory parameters in the first 24 h of ICU stay such as CRP, PCT, ferritin, IL-6, lymphocyte-, leukocyte-, neutrophil-, and granulocyte count, D-dimer, history of the comorbidities like HT, DM, COPD [with differentiating those who used inhalational corticosteroids prior to ICU admission], chronic kidney disease (CKD). An upper limit was applied for CRP, PCT, ferritin, IL-6 and D-dimer based on the first limit of laboratory measurements.

These interventions did not cause any differences in any of the statistical significances, however, created an opportunity to design clearer figures and tables with less distortion by outlier values. The database from the University of Debrecen included the following data: gender, age, date of testing, days spent in ICu, chest CTSS, P/F ratio, laboratory parameters and comorbidities.

The latter half of our study included a ML algorithm-based analysis called RF. The primary analysis of 503 patients in Pécs proved to be a small number of cases and low reliability values were obtained when using complex statistical methods, therefore we decided to collaborate. From the ICUs of the University of Debrecen, we managed to enroll an additional 124 patients. For these patients, no viral sequencing data were available, so in addition to the three main variants (Alpha, Delta, Omicron), we included 22 patients with B.1.160 variants from our database from patients in Pécs, resulting in a total of 649 patients. A graph showing the steps of the analysis is shown in Figure 1.

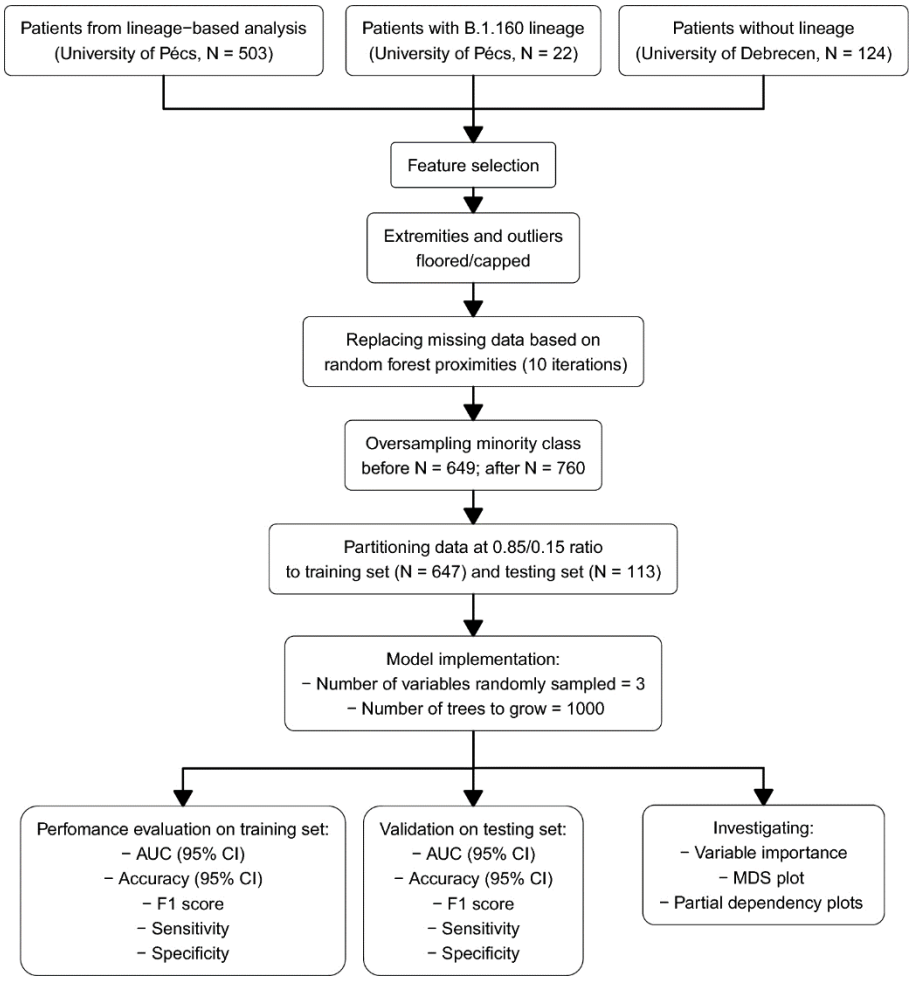

Figure 1: Workflow chart of the RF analysis

In our RF model, we used nine parameters: Age, P/F ratio, chest CTSS, Days at ICU, D-dimer, PCT, CRP, Leukocyte count, Lymphocyte count as predictive features and ICU mortality as a binary outcome (Deceased, Survivor). We selected predictors based on their potential importance and to account for multicollinearity, excluding those which are highly correlated with one another. Since several parameters experienced numbers of missing data points, for example, up to 16% in the case of the Horowitz index, we used data imputation prior to model fitting. RF approximation matrices (which measures the similarity of data points) are used to fill in missing data. Given that the mortality outcome was not balanced (59% died - 41% survived), a balancing step had to be performed. The balancing, i.e. the oversampling of the minority class (survivors), was necessary to achieve a ratio of 0.5 (50% survived, 50% deceased), as this allows efficient learning and more accurate estimation of the model. Without oversampling, the model would be dominated by deceased cases, if the model learns from this, a bias would be introduced, and the model would subsequently tend to overestimate the probability of mortality on a new (previously not seen) unknown group by the model. More details on imputation and balancing are available in the dissertation. After oversampling, we acquired 760 data points. Then the data were partitioned randomly to a training and testing set in 0.85/0.15 ratio (training data, N = 647, testing data, N = 113). The training dataset was used to fit the RF model. After model fitting, we investigated the RF model performance on the test dataset, defining the most widely used metrics: accuracy with 95% confidence intervals (CI), sensitivity, specificity, and area under the curve (AUC). To calculate these values, we used the model predictions to predict new data of the training and testing datasets. We investigated how well the model performed on supplied training data, and also how it performed regarding previously unseen data. Furthermore, we calculated 95% confidence intervals with an exact method for the AUC and accuracy metrics, therefore, we have far more than merely a point estimate. Based on the mean decrease in the Gini score and upon model accuracy, we determined the order of variable importance as well.

The technical details of whole genome sequencing (WGS) and the ethical implications are described in the dissertation.

All statistical calculations were performed in R Statistics version 4.3.2 (R Foundation for Statistical Computing, Vienna, Austria). A p-value < 0.05 was defined as a two-tailed level of significance.

IV. Results

IV/1. Dynamic, interactive epidemiological map

The data collection covers the timespan between 2020.08.19 and 2022.02.13 and includes 271,849 COVID-19 RT-qPCR test related data. Broken down by individuals, this is 158,036 persons belonging to Baranya, Somogy and Tolna counties. The overall median age was 44 years (IQR 27-61), with a male-female split of 43.3% / 56.7%. The percentage breakdown of tested individuals by county was as follows: Baranya - 55.26%, Somogy - 22.05%, Tolna - 22.69%. The highest percentage of positive PCR tests was found in Somogy (26.1%), Tolna (23.3%), while the lowest percentage was found in Baranya County (16.8%). The distribution of symptomatic individuals was also highest in Somogy County (72.2%), followed by Tolna County (62.9%) and lowest in Baranya County (47.3%). The total number of tests, positive test results, asymptomatic and symptomatic cases per 656 municipalities in the three counties can be found on covid-pte.vercel.app/dashboard. Two data to highlight: the average age of the tested individuals in Pécs was 26.3 years until 31 August 2020. The average age in Cserénfa until 31 December 2020 was 80.1 years, and the average cycle threshold was 24.54.

Sequencing results

During the time interval studied, 77 different SARS-CoV-2 lineages were identified from 2,975 individuals, the vast majority of which can be classified as European, Alpha, Delta and Omicron variants. The exact lineages by time interval are reported in the dissertation. If the sequenced lineages are grouped according to the main viral variants, four groups are obtained: European, Alpha, Delta, and Omicron groups. There was a significant difference ($p<0.001$) in median age between groups: European - 65 years, Alpha - 54 years, Delta - 46 years, Omicron - 40 years. When age was divided into 5 different subgroups, a significant difference ($p<0.001$) was also observed among the groups. There was no significant difference ($p=0.249$) in the sex distribution. There

was also a significant difference ($p<0.001$) in the presence of symptoms between the four groups of viral variants. For the median cycle threshold values, a significant difference ($p=0.003$) was also observed between groups: lowest for Delta (25.3), Alpha (25.5), followed by European (26.0) and Omicron (26.2).

IV/2. Diagnostic performance of Panbio RAT

A total number of 5,136 samples tested parallel for SARS-CoV-2 Panbio RADT and SARS-CoV-2 RT-qPCR were collected from 4,440 individuals who were admitted in the clinical departments of the University of Pécs, Clinical Center. The tested individuals were between 0 and 101 years old (median age: 53 years, IQR 30–72 years). The female/male ratio was 57.2%/42.8%. The median Ct values were significantly lower in the symptomatic group when compared with the asymptomatic group (28.2 vs. 35.0, respectively $p < 0.001$). To perform a stratified statistical analysis, 696 parallel samples were excluded to avoid distortion, which were repetitive tests of patients who were aligned to the follow-up of the SARS-CoV-2 RT-qPCR positive cases. Out of 4,440 paired tests, 609 samples tested positive using RT-qPCR, resulting in a prevalence of 13.7%. Panbio detected 251 (5.7%) positive tested samples. In this calculation method, overall sensitivity was 41.2%, overall specificity was 99.7%. Positive predictive value (PPV) was 95.1%, negative predictive value (NPV) was 91.4%.

	Panbio positive	Panbio negative	
PCR positive	251 (5.7%)	358 (8.1%)	Sensitivity: 41.2% (95% CI 37.4-45.2%)
PCR negative	13 (0.3%)	3,818 (86.0%)	Specificity: 99,7% (95% CI 99.4-99.8%)
	PPV: 95.1% (95% CI 91.8-97.1%)	NPV: 91.4% (95% CI 90.5-92.2%)	

Table 1: Analytic performance of Panbio, repetitive tests (696) excluded

When we examined the Panbio RAT sensitivity among Cycle threshold subgroups, it was 91.2%, in the group of Ct values ≤ 20, 68.6% within the Ct range of 20–25, 47.9% in the group of Ct values between 25 and 30, and 12.6% in the group of Ct values between 30 and 35.

We compiled demographic and clinical data from 80 individuals who succumbed due to complications related to SARS-CoV-2 and compared it with the generally tested population. There was a remarkable difference in gender distribution of the tested cases among the general population who were SARS-CoV-2 suspected, in favor of females: 42.5 vs. 57.5% (male:female ratio). The median age difference was also significant ($p < 0.001$): 52 (IQR 30–71) in the general population vs. 78 (IQR 70–87) among the fatal cases.

IV/3. Study of clinical, laboratory and molecular genetic parameters associated with adverse outcomes of SARS-CoV-2 infection

317 out of 503 ICU patients succumbed while in the ICU. Mortality was 65.5% in the Alpha group (127/194 patients), 66.1% in the Delta group (152/230 patients) and 48.1% in the Omicron group (38/79 patients). Kaplan–Meier survival curve is one of the best options to measure the fraction of subjects living for a certain amount of time following treatment, which in our case, begins with ICU admission.

Figure 1, panel (a) shows the survival curves by lineage. A 50% survival probability was reached on Day 15 in the case of Alpha-, Day 14 by Delta-, and Day 13 by Omicron patients. There was no significant difference among the lineages ($p = 0.95$). We also used Kaplan–Meier to compare survival probability among the three age groups (< 50, 50–65, > 65, see Fig. 1, panel (b). The youngest group of ICU patients aged < 50 years old reached a 50% survival probability on Day 22, while the middle group of patients aged between 50 and 65 years old reached a 50% survival probability on Day 16. The oldest group with age > 65 reached a 50% survival probability on Day 12. There was a very highly significant difference ($p < 0.0001$) among the age groups.

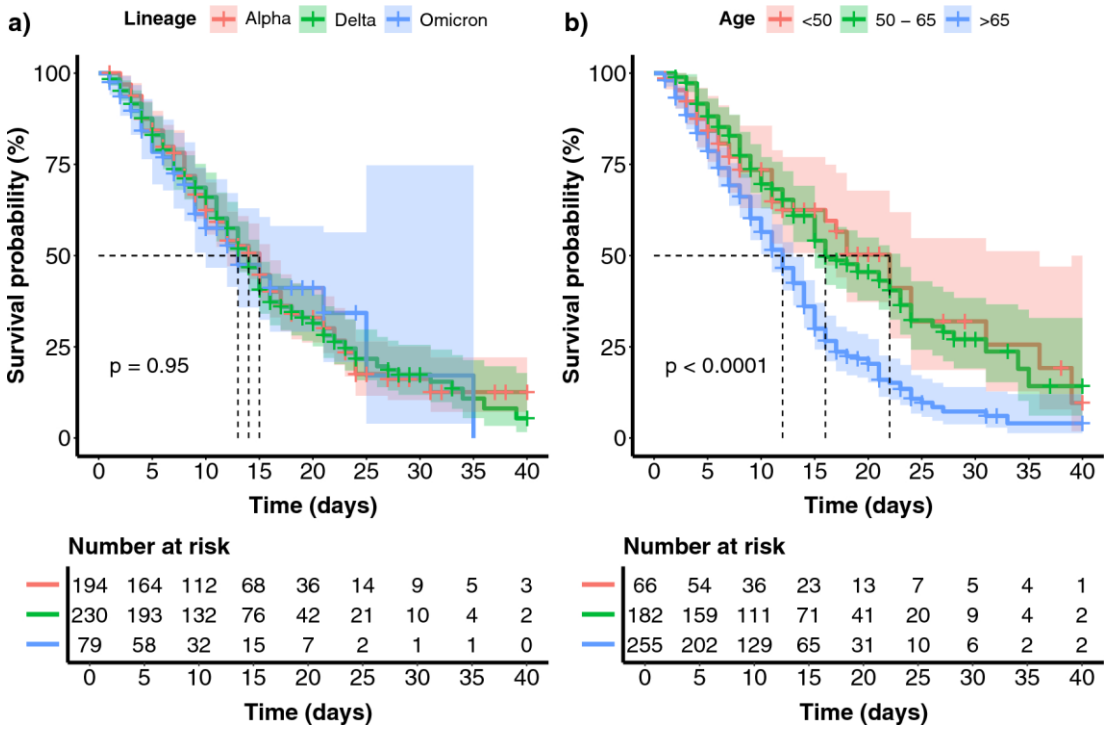

Figure 2: Kaplan-Meier survival probability curve

a) Survival probability among Alpha, Delta and Omicron VOCs

b) Survival probability among the age groups < 50, 50-65 and > 65 years old

Vaccination data

In regard to vaccination, 339 out of 503 patients had no vaccination history, and among these unvaccinated patients, 222 patients (65.5%) succumbed. 35 patients received only one vaccination dose, in which 24 (68.6%) patients in this group soon after they expired. Patients with two vaccination doses following 14 days from the second vaccination dose were considered fully vaccinated; 82 patients were fully vaccinated, and 53 (64.6%) patients in this group expired. In reference to those patients who received a booster dose, 47 patients received 3 doses of vaccination, in which only 18 (38.3%) patients expired in this group. Patients were considered fully vaccinated 14 days following the second dose of vaccination, whereas fully vaccinated patients who received at least two doses of vaccination and/or at least one booster dose within a six-month window post-vaccination from the first booster dose were considered protected patients. Out of 503 patients, only 55 were deemed protected, in which 27 (49.1%) patients expired in this group. In consideration of the 437 non-protected patients, 281 (64.3%) patients succumbed. It is noteworthy in highlighting, the remaining 11 Delta-infected patients were fully vaccinated, however, sufficient data in

terms of vaccination dates were missing; 9 of these 11 patients were deceased. When we performed Pearson's Chi-squared test, it revealed the protective effect of vaccination significantly reduces mortality among protected patients with a p-value of 0.028, see Table 2.

Variables	**Vaccinated**				Total	p-value	**Protected**		Total	p-value
	Not (0)	**Partially (1)**	**Fully (2)**	**Booster (3)**			No	Yes		
Mortality						**0.003**[1]				**0.028**[1]
Survived	117 (35%)	11 (31%)	29 (35%)	29 (62%)	186 (37%)		156 (36%)	28 (51%)	184 (37%)	
Deceased	222 (65%)	24 (69%)	53 (65%)	18 (38%)	317 (63%)		281 (64%)	27 (49%)	308 (63%)	
Total, n (%)	339 (100%)	35 (100%)	82 (100%)	47 (100%)	503 (100%)		437 (100%)	55 (100%)	492 (100%)	
[1] Pearson's Chi-squared test										

Table 2: Mortality data among non-vaccinated and vaccinated ICU patients

Random forest

The RF analysis contains 649 ICU patients' data: 503 from the original database, 22 patients infected with the lineage B.1.160, and 124 patients from the University of Debrecen. Following imputation and balancing, the model tallies 760 patients. The test model performance had an accuracy of 0.814 (95% CI 0.73–0.881), with a p-value of < 0.0001. Sensitivity was 0.825, specificity was 0.804. Additional performance data and the Receiver Operation Characteristics (ROC) curve are illustrated in Figure 3.

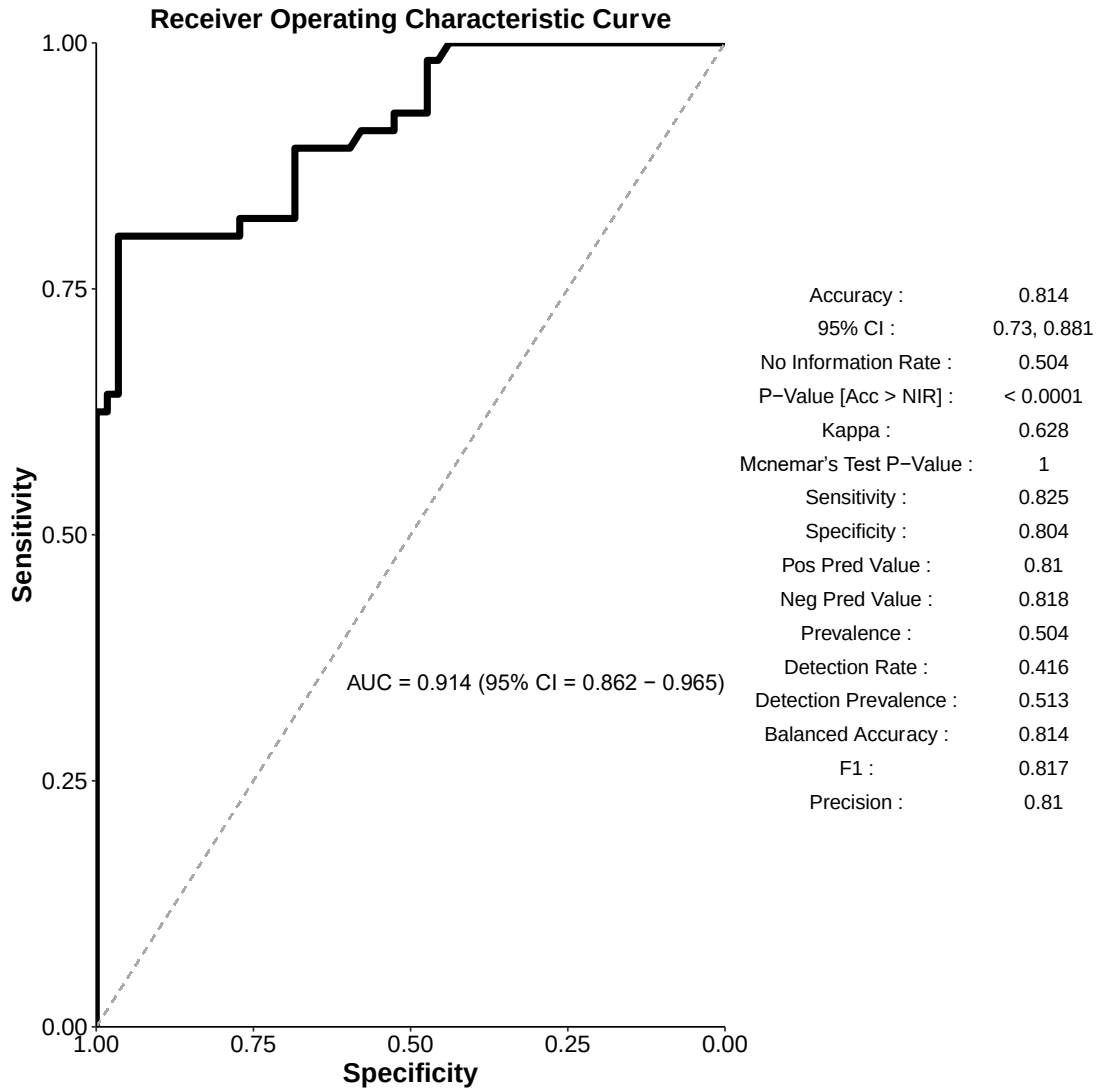

Figure 3: Test model performance metrics with the ROC curve

MeanDecreaseAccuracy (MDA) and MeanDecreaseGini (MDG) values of the key parameters are shown on Figure 4.

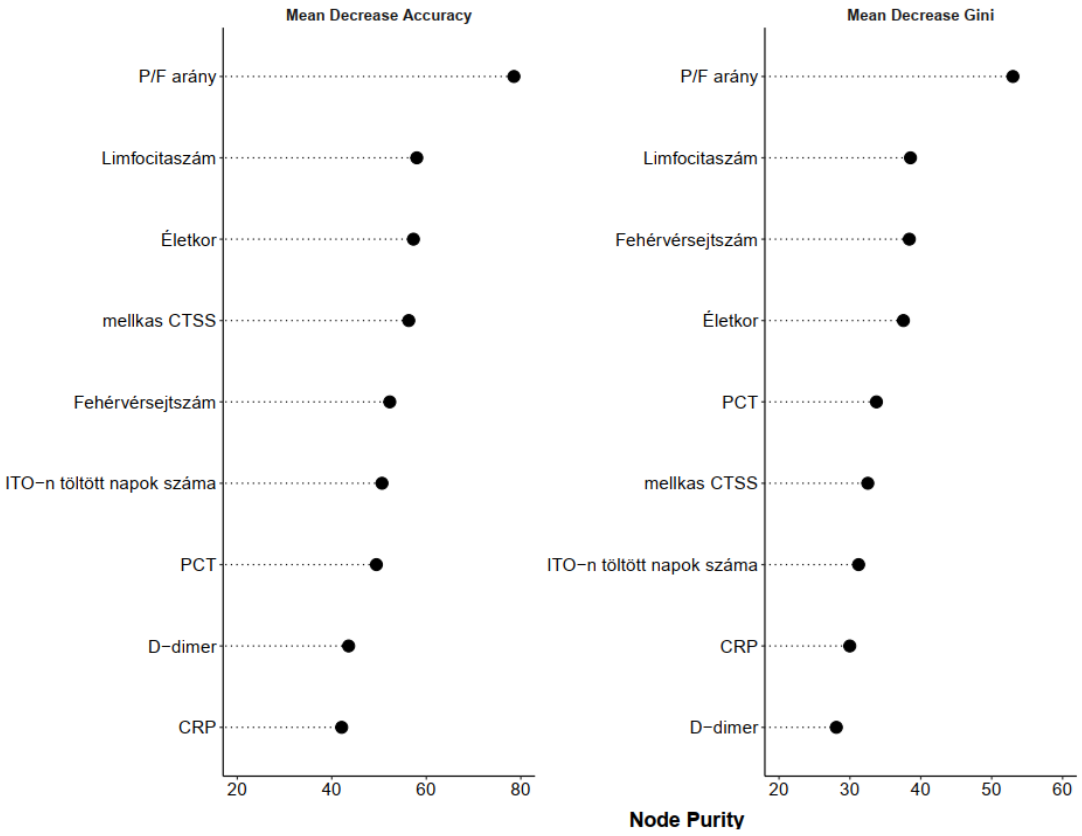

Figure 4: MeanDecreaseAccuracy (MDA) and MeanDecreaseGini (MDG) values of the key parameters

Detailed statistics on clinical and laboratory parameters, broken down into variants and survivor/death groups, are available in the dissertation.

V. Discussion

V/1. Dynamic, interactive epidemiological map

If we look at the data for Pécs before 2020.09.02., the average age in the first two weeks of the period under review was only 25.9 years. The reason is that the very first tests came from different freshman camps of PTE, where we detected the first positive cases. Thereafter, the average age increased gradually (peak 50.9 years), and with the appearance of the Omicron variant with higher contagiosity, the average age of infected started to decrease again, reaching 48.4 years at the end of our study. Another noteworthy phenomenon is the identification of epidemiological clusters with specific demographic data on our map, e.g. according to the nursing homes in each municipality. Cserénfa is a tiny settlement with only 217 inhabitants in Somogy County, where we identified an outstanding number of positive cases. The first tests

were received in November 2020: 26 positive tests were identified in one month, with an average age of 80.1 years and an average cycle threshold of 24.54, indicating a considerably high viral load. Following the emergence of SARS-CoV-2 in Hungary, nursing homes and chronic care facilities were well identified as epidemiological hotspots during the short upsurge phase of the wild-type dominated epidemic period. Baranya County had a significantly lower prevalence of symptomatic patients (47.3%) than Somogy (72.2%) or Tolna (62.9%) counties. The proportion of positive cases also shows a similar trend: 16.8% in Baranya County, compared to 26.1% and 23.3% in Somogy and Tolna counties respectively. The reason for these differences is most likely due to the higher screening capacity of Baranya County (and mainly of the PTE KK), which included health professionals and students, while in the other two counties testing was more symptom oriented. Regarding the incidence of cases, the timeline-like fluctuation of cases observed in our region during our study period perfectly matched the data set representing the whole of Hungary, and some trends, such as the drastic decrease in the number of cases in summer, which was previously suggested by several studies (Aboubakr et al., 2021)(Chen et al., 2021), were also observed.

Reflecting on molecular epidemiology, the 77 sublineages detected by us are not surprising, a Slovenian study (Janezic et al., 2023), which genomically analyzed the first 9 months of 2021, identified 64 sublineages, and the variations and trends (including the low number of cases in summer) that appeared in the study in a timeline are similar to the changes observed in our study.

V/2. Diagnostic performance of Panbio RAT

Panbio RAT was reported to maintain a high specificity (between 94.9 and 100%) in preliminary clinical studies (Fenollar et al., 2021)(Gremmels et al., 2021). A study from Heidelberg, Germany demonstrated a sensitivity of 95.8% in Ct values <25 and within seven days from symptom onset (Krüger et al., 2021). In larger study populations, Panbio sensitivity was between 33.3% (Masiá et al., 2021) and 55.3% (Landaas et al., 2021) in asymptomatic patients. A German study with more than 1,000 participants described an overall sensitivity of 46.7% (Wagenhäuser et al., 2021), while an Italian study with 4,167 participants reported a sensitivity of 66.8% (Treggiari et al., 2022). Our study began with a review of 5,136 cases in which we found SARS-CoV-2 Panbio RADT overall sensitivity to be low, at 36.1%. Test sensitivity improved

to 41.2%, when repetitive follow-up tests were excluded from the analysis, which is primarily due to the exclusion of samples with low viral load close to the maintained cut-off level and the lowest detection limit of the qPCR.

According to a study from Cologne, Germany (Platten et al., 2021), 52.6% of positive cases with Ct values > 28 were undetected by RAD tests. Our findings are consistent with the above-mentioned study: the 80 patients who succumbed due to complications of SARS-CoV-2 had a median Ct value of 27.0, with a Panbio sensitivity of 47.5%. Results of our clinical study highlight the universal observation associated with SARS-CoV-2 RT-qPCR: it is the most reliable tool in the detection of active SARS-CoV-2 infection. Although SARS-CoV-2 RADT offers several advantages over SARS-CoV-2 RT-qPCR, even in clinical settings due to its point-of-care testing (POCT) administration and rapid turnaround time, these tests are less sensitive or at critically low prevalence rate of the infection and can be considered unsatisfactory regarding accurate testing and consequential diagnosis.

In the combined SARS-CoV-2 Panbio RADT and RT-qPCR tested population, 80 fatal cases were observed during our study period. Statistical evaluation of the deceased population group identified significant differences compared to the later recovering general population according to gender, age and presence of the symptoms during the first testing and PCR Ct stratification distribution. Older age, male sex, clinically symptomatic status and lower Ct range are all significantly correlated to disease fatality. However, we emphasize, six of the patients (7.5%) had no clinical symptoms during the first test and SARS-CoV-2 Panbio RADT was negative in 38 individuals, 47.5% of the fatal cases. These numbers suggest rapid antigen testing should not be the sole test administered to populations at high risk of developing severe disease.

We had no information regarding the onset of symptoms, which should be taken into consideration when comparing it with diagnostic performance. There was no possibility to repeat the RAD or RT-qPCR tests from the same samples due to continuous high daily activity, which leaves open the possibility of human error (e.g., RAD test evaluation beyond the recommended timeframe), despite being performed and evaluated strictly by healthcare professionals.

V/3. Study of clinical, laboratory and molecular genetic parameters associated with adverse outcomes of SARS-CoV-2 infection

Age is found to be the highest risk factor in COVID-19 infection, likely influencing all molecular mechanisms from immune responses, mitochondrial functions, endoplasmic reticulum transport mechanisms and protein folding, oxidative stress disruption, receptor activation of the ACE II and Toll-like receptors, transcription factors and cell signaling pathways (Chatterjee et al., 2023).

ICU mortality spanned a vast frontier (28.8-75.9%) among studies examined in the literature. Our findings reveal an ICU hard mortality of 63.0%, which aligns with the only Hungarian COVID-19 ICU study published to this point (Nagy et al., 2023). Based on our timeline and lineage distribution, Omicron sublineages had approximately as much time (January 2022 to November 2022) when compared to Alpha and Delta combined (January 2021 to December 2021). Nevertheless, patients admitted to the ICU because of Omicron only represent 15.7% of total cases, which implies ICU admission was significantly lower in the case of Omicron. This can potentially be explained by the effects of protection due to vaccination and milder lineage combined. The Omicron variant is less severe; however, once admitted to the ICU, Omicron still seemingly proves to be a potentially deadly variant since there is no significant difference among the lineages in terms of survival rate, see Kaplan–Meier survival curves by lineage in Fig. 1, panel a).

In regard to the Hungarian population, vaccination strategies were initially introduced in December 2020. Vaccine uptake of the primary course counts up to 63.2% of the total population, with 39.8% taking up the first booster, and 4.3% in the second booster up through 4 April 2023, in the general population. Among the population admitted to the ICU in the time period of our study, 35% were unvaccinated and of those who later survived when compared to the 65% of unvaccinated patients who succumbed in COVID-19. In calculating only the full primary course, meaning the vaccinated and boosted individuals within the protective time period, vaccines are considered to have effective protection (minimum 14 days—maximum 6 months), in which we observed a significant mortality-reducing effect of the vaccination. Based on our results, we recommend those individuals over age 65 and comorbid populations will benefit from vaccination on a priority basis when compared to individuals without these conditions.

Based on the random forest analysis, we were able to demonstrate the predictive capacity regarding the machine-learning approach in assessing ICU patient lethality due to COVID-19. The analysis - incorporating data from 649 ICU patients - achieved an accuracy of 86.24% in training and 81.4% in testing phases, underscoring its robustness. Key predictors included the P/F ratio, lymphocyte count, and chest CTSS, among others, indicating the relevance of respiratory status, immune response and lung involvement in determining patient outcomes. Another key predictor of our model was the age of ICU patients, which was also highlighted in an Italian study (Lorenzoni et al., 2021) as the leading predictor in their models.

In recent years, various MDL methods have been implemented to predict COVID-19 mortality, even in ICU settings. According to Shen (et al., 2023), the RF model has the best performance in predicting the risk of death in hospitalized patients with COVID-19. From a technical point-of-view, the most commonly used MDL methods were random forest, logistic regression, and decision tree (commonly gradient boosted, such as XGBoost)(Viderman et al., 2024). Several studies focused on comparing the performance of multiple MDL methods, while others focused more on the clinical point-of-view, exploring different demographic- and/or clinical- and laboratory parameters, and comorbidities in their models (Elhazmi et al., 2022)(D. Li et al., 2020). Shi (et al., 2020) revealed that RF showed the best performance out of three machine learning models to predict COVID-19 mortality, with the top three important variables being mean arterial pressure, age and PCT. Sakagianni (et al., 2023) also found the RF as the best outcome predictor in COVID-19 ICU patients, with urea, age, hemoglobin, CRP, platelet count and lymphocyte count as the top six important variables. Jamshidi (et al., 2022) compared several machine learning algorithms to predict mortality in day 0 ICU patients, with 15 factors, mostly laboratory parameters. Random forest outperformed other models and had a superior efficiency, parameters giving the most information on the probability of each patient's death were albumin, urea, red blood cell distribution width and age. Another Iranian study (Najafi-Vosough et al., 2023) found age as the most important variable for predicting mortality in their RF analysis.

Unlike many conventional models, which may rely on a narrower set of variables, the RF approach allows for the integration of a wide range of clinical parameters, enhancing its predictive capability (Ovcharenko et al., 2023), (C. Zhan et al., 2021), (Zhao et al., 2022). The practical significance of the recent study lies in its potential to

support clinical decision-making by offering a model that combines high predictive accuracy with practical applicability, in which the effects of vaccination and newer VOCs are taken into consideration in a well-defined ICU population with clear exclusion criteria. In providing an accurate prediction of patient outcomes, ICU specialists can customize interventions more effectively and improve patient care and resource allocation. To cite an example, patients identified as high-risk can be prioritized for more aggressive treatment or monitoring. A further advantage regarding the RF is the model can manage non-linear relationships and interactions between variables, enhancing its predictive capability across diverse clinical scenarios. Moreover, RF is known for its resistance to overfitting, specifically with the use of multiple trees, assuring its reliability for practical applications.

Since we only had Hungarian patients' data for model training and did not validate our result on a dataset from a different origin, it is unknown whether it is generalizable to other countries or could be only applied to the Hungarian patient population. However, we chose widely used clinical parameters (such as age, P/F ratio, and lymphocyte count). Therefore, we believe that our predictions could be useful for other researchers and clinicians as well. Another limiting factor is the model's interpretability. While RF offers high accuracy, it is inherently more complex and less interpretable than simpler models, which may pose challenges for clinical communication, even though we tried to enhance its transparency by utilizing feature importance. The model's performance heavily relies on the quality and completeness of input data, making it sensitive to any biases or inaccuracies present, which could be affected by the retrospective nature of the study. These strengths and limitations may underscore the potential of the random forest analysis in clinical applications while highlighting areas for improvement and careful consideration in future research and implementation.

In conclusion, when considering the application of a machine learning algorithm, our study promoted a comprehension of the mechanisms and hierarchically estimated the risk-modifying effects of demographical factors and pathophysiological and pathobiochemical parameters among patients coping with the severe course of SARS-CoV-2 infection. Advanced bioinformatical analysis regarding clinical data can potentially enable clinicians to customize guidelines and to develop care strategies and treatment alternatives, suited for the most vulnerable populations in intensive care and potentially guiding more personalized, timely interventions. Future studies can explore

the integration of additional variables, such as genetic markers or detailed clinical history, in larger, externally validated, multi-center studies with patients from diverse nations and races to refine the model's predictive accuracy. Moreover, the application of the model in prospective clinical trials will provide valuable insights into its real-world effectiveness and impact upon patient outcomes.

VI. Conclusion

The results obtained from the epidemiological map analysis and the Panbio rapid antigen test study provide a comprehensive picture of the epidemiological and diagnostic challenges during the COVID-19 outbreak, as well as the effectiveness and limitations of the methods used. The timely identification of epidemiological clusters, the settlements with high infection rates and the underlying institutional infrastructure underlines the importance of both regional and institutional level epidemiological control in a flexible and managed way. In Baranya County, where Pécs is the main city, the specialized molecular diagnostic units of the University of Pécs Clinical Centre allowed for a higher screening capacity, resulting in an overall lower rate of positive cases and symptomatic cases. Increasing screening capacity proved effective in reducing the spread of the epidemic and enhancing the efficiency of epidemic management. Our data show that the fluctuation in positive case rates in the South Transdanubia region was in line with national trends, confirming that the local epidemiological situation is closely linked to national trends.

The diagnostic performance of the rapid antigen test for mass screening was found to be low due to the low sensitivity of the test. This indicates that the use of the test may be limited in cases with low population prevalence of infection, early stages of infection and detection of cases with low viral shedding. However, it is an appropriate and rapid alternative for samples with high virus concentrations and high prevalence of infection in the community phase of the epidemic. Genetic material of SARS-CoV-2 was detectable for up to 35 days, which was not strongly correlated with patient infectivity. This suggests that PCR positivity does not always indicate infectivity, and diagnostic strategies should take this factor into account.

Older age, male sex, clinical symptoms and high viral shedding were significantly correlated with mortality. This highlights the importance of special surveillance of

older populations with chronic disease during the epidemic. Our results also confirm that vaccination significantly reduced mortality, especially among older patients with co-morbidities. This supports the importance of vaccination in preventing severe disease outcomes. Results of RF analysis showed that machine learning methods are effective tools for predicting severe cases of COVID-19. This approach has the potential to support clinical decision making, particularly in the identification and management of high-risk patients.

Based on the above, a number of guidelines and suggestions for the development of future epidemiological and diagnostic strategies can be formulated: epidemiological data clearly show that increasing screening capacity is key to effectively managing the epidemic. A high screening capacity can help to identify asymptomatic cases and control the spread of the epidemic. It is reasonable to design epidemiological measures taking regional differences into account. In counties with lower screening capacity, increased attention should be paid to expanding screening and strengthening prevention measures. Refinement of testing protocols and the use of combined diagnostic approaches can help to achieve more accurate results. The integrated use of machine learning methods in the analysis of epidemiological data and the prediction of serious cases is a promising trend. Future research and development should also focus on further refining these models and incorporating them into clinical practice.

Overall, the experience and research findings from the management of the COVID-19 epidemic highlight the importance of an adaptive and dynamic approach in the development of epidemiological strategies, diagnostic methods and predictive decision support tools. Building on our current experience, there is potential to develop even more effective tools and strategies for controlling epidemic processes in the future.

VII. New results

Dynamic, interactive epidemiologic map

a) Representation of a specific map for the Southern Transdanubia region, based on diagnostic and demographic data, which, if sufficiently detailed data are available, can be used in the future to represent other epidemiological/other public health updates.

b) Timeline presentation of daily epidemiological data (positive case counts, cycle threshold values, total test count, average age) per population.

c) Presentation of regional virus variants over the study period, broken down by municipality and displayed on a map.

Diagnostic performance of Panbio RAT

a) Evaluation of the diagnostic performance of Panbio RAT in clinical settings for 5,136 cases enrolled, which at the time of publication was the largest clinical setting of the COVID-19 Panbio RAT study internationally.

b) To make recommendations for combined (RAT + PCR) testing based on the performance evaluation of the Panbio test.

Study of clinical, laboratory and molecular genetic parameters associated with adverse outcomes of SARS-CoV-2 infection

a) Comparison of the first laboratory parameters within 24 hours of ICU admission among the 3 COVID-19 viral variants (Alpha, Delta, Omicron), which is a novelty in international comparison.

b) An analysis of the association between COVID-19 vaccination status and mortality in a Hungarian intensive care unit population.

c) Identification of the most important factors associated with lethality in ICUs.

d) Development of a Random Forest algorithm based on laboratory and clinical parameters most strongly correlated with lethality, aimed at supporting clinical decision-making.

VIII. Acknowledgements

I would like to take this opportunity to express my gratitude to everyone who contributed to the completion of this dissertation and supported me throughout the past years. First and foremost, I am deeply grateful to my supervisor, Dr. Katalin Gombos, for her trust in me and her continuous support. Her outstanding professional guidance and personal assistance will serve as an example for me throughout my career. I had many opportunities to learn from her and alongside her. I owe special thanks to Alex Váradi; without his background work in biostatistics and informatics, neither this dissertation nor our scientific publications would have been possible. I am also deeply grateful to my friend and fellow PhD student, Dr. Kristóf Filipánits, with whom I have worked extensively since our final year of medical school, and who has provided me with numerous great ideas and advice during my PhD studies. I would also like to thank my childhood friend, Róbert Wágner (PCSUNIQ Ltd.), for our collaborative work, without whom the interactive map of the Southern Transdanubia region would not have been created. I am thankful to the analysts at the Institute of Laboratory Medicine and my fellow PhD students for their help and support. I also extend my gratitude to the staff of the PTE Clinical Centre's Institute of Medical Microbiology and Immunology who participated in SARS-CoV-2 diagnostics. From the Molecular Medicine Research Group at the Szentágothai Research Centre, I would like to thank my colleague Daryan Jamal Mohammed for his hard work, assistance, and valuable discussions related to our research in the intensive care unit. I would like to thank Dr. Krisztián Jáksó for his professional help in the intensive care unit, and the collaborators from the Szentágothai Research Centre Genomics and Bioinformatics Core Facility and also from the University of Debrecen for their collaboration. Finally, I would like to thank my family and close friends who have been by my side all the way, without them I would not have made it this far.

The research and author of this dissertation was sponsored by:

- Thematic Excellence Program (TKP2021-EGA-13) project
- PTE Research Grant (009_2023_PTE_RK/25)
- New National Excellence Program (ÚNKP-23-3-II-2085)

IX. List of publications

Articles related to this thesis

1. **Hamar Á**, Filipánits K, Váradi A, Váradi-Rácz R, Gellén HO, Futács K, Urbán P, Kovacs GL and Gombos K (2022). Diagnostic accuracy of SARS-CoV-2 Panbio™ rapid antigen diagnostic tests in a 4,440-case clinical follow-up. Front. Med. 9:908127. doi: 10.3389/fmed.2022.908127
https://www.frontiersin.org/articles/10.3389/fmed.2022.908127/full
Frontiers in Medicine – Q1
Impact Factor: 3.9
Number of independent citations: 6

2. **Hamar Á.**, Mohammed D., Váradi A., Herczeg R., Balázsfalvi N., Fülesdi B., László I., Gömöri L., Gergely PA., Kovács LG., Jáksó K., Gombos K.
COVID-19 mortality prediction in Hungarian ICU settings implementing random forest algorithm. Sci Rep 14, 11941 (2024). https://doi.org/10.1038/s41598-024-62791-9
Nature Scientific Reports – D1, Q1
Impact Factor: 3.8
Number of independent citations: 2

Cumulative Impact Factor: 7.7
Cumulative Q value: Q1: 100%, D1: 50%
Cumulative independent citations: 8

Conference presentations and posters related to my thesis:

1. **Hamar Á**, Filipánits K, Glavatity A, Wágner R, Gombos K. Covid-19 pandemic in the Southern Transdanubia region depicted using a dynamic interactive map. X. Jubileumi Interdiszciplináris Doktorandusz Konferencia; Pécs, 2021.11.12-13. (presentation)

2. **Hamar Á**, Herczeg R, Wágner R, Filipánits K, Gombos K. SARS-CoV-2 PCR diagnostics and subsequent epidemiologic analysis implementing a dynamic interactive map depicting the southern Transdanubia region in Hungary. 32nd

European Congress of Clinical Microbiology and Infectious Diseases, Lisszabon, Portugália, 2022.04.23-26. (poster)

3. **Hamar Á.**, Váradi A., Wágner R., Filipánits K., Gombos K. SARS-CoV-2 PCR diagnosztika és epidemiológiai analízis a dél-dunántúli régióban egy dinamikus, interaktív térkép segítségével. MOLSZE XVII. Nagygyűlése, Budapest, 2022.08.26-27. (presentation)

4. **Hamar Á.**, Váradi A., Wágner R., Filipánits K., Gombos K. SARS-CoV-2 PCR diagnosztika és epidemiológiai analízis a dél-dunántúli régióban egy dinamikus, interaktív térkép segítségével. I. Romhányi György Szakkollégium Konferencia, Pécs, 2022.11.04-06. (presentation)

5. **Hamar Á.**, Váradi A., Wágner R., Gombos K. SARS-CoV-2 PCR diagnostics and epidemiologic study depicted on a regional dynamic map. 19th International Medical Ph.D. Conference, Hradec Kralove, Csehország, 2022.11.23-25. (presentation)

6. **Hamar Á.**, Mohammed D., Váradi A., Herczeg R., Jáksó K., Balázsfalvi N., Gombos K. COVID-19 letalitás predikció az intenzív osztályon: Random Forest algoritmus alkalmazása klinikai és laboratóriumi paraméterekkel. II. Romhányi György Szakkollégium Konferencia, Pécs, 2024.03.01-03. (presentation)

X. References

Aboubakr, H. A., Sharafeldin, T. A., & Goyal, S. M. (2021). Stability of SARS-CoV-2 and other coronaviruses in the environment and on common touch surfaces and the influence of climatic conditions: A review. Transboundary and Emerging Diseases, 68(2), 296–312. https://doi.org/10.1111/tbed.13707

Breiman, L. (2001). Random Forests. *Machine Learning*, *45*(1), 5–32. https://doi.org/10.1023/A:1010933404324

Chatterjee, S., Nalla, L. V., Sharma, M., Sharma, N., Singh, A. A., Malim, F. M., Ghatage, M., Mukarram, M., Pawar, A., Parihar, N., Arya, N., & Khairnar, A. (2023). Association of COVID-19 with Comorbidities: An Update. ACS Pharmacology & Translational Science, 6(3), 334–354. https://doi.org/10.1021/acsptsci.2c00181

Chen, S., Prettner, K., Kuhn, M., Geldsetzer, P., Wang, C., Bärnighausen, T., & Bloom, D. E. (2021). Climate and the spread of COVID-19. Scientific Reports, 11(1), 9042. https://doi.org/10.1038/s41598-021-87692-z

Elhazmi, A., Al-Omari, A., Sallam, H., Mufti, H. N., Rabie, A. A., Alshahrani, M., Mady, A., Alghamdi, A., Altalaq, A., Azzam, M. H., Sindi, A., Kharaba, A., Al-Aseri, Z. A., Almekhlafi, G. A., Tashkandi, W., Alajmi, S. A., Faqihi, F., Alharthy, A., Al-Tawfiq, J. A., … Arabi, Y. M. (2022).

Machine learning decision tree algorithm role for predicting mortality in critically ill adult COVID-19 patients admitted to the ICU. Journal of Infection and Public Health, 15(7), 826–834. https://doi.org/10.1016/j.jiph.2022.06.008

Fenollar, F., Bouam, A., Ballouche, M., Fuster, L., Prudent, E., Colson, P., Tissot-Dupont, H., Million, M., Drancourt, M., Raoult, D., & Fournier, P.-E. (2021). Evaluation of the Panbio COVID-19 Rapid Antigen Detection Test Device for the Screening of Patients with COVID-19. Journal of Clinical Microbiology, 59(2). https://doi.org/10.1128/JCM.02589-20

Gremmels, H., Winkel, B. M. F., Schuurman, R., Rosingh, A., Rigter, N. A. M., Rodriguez, O., Ubijaan, J., Wensing, A. M. J., Bonten, M. J. M., & Hofstra, L. M. (2021). Real-life validation of the PanbioTM COVID-19 antigen rapid test (Abbott) in community-dwelling subjects with symptoms of potential SARS-CoV-2 infection. EClinicalMedicine, 31, 100677. https://doi.org/10.1016/j.eclinm.2020.100677

Jamshidi, E., Asgary, A., Tavakoli, N., Zali, A., Setareh, S., Esmaily, H., Jamaldini, S. H., Daaee, A., Babajani, A., Sendani Kashi, M. A., Jamshidi, M., Jamal Rahi, S., & Mansouri, N. (2022). Using Machine Learning to Predict Mortality for COVID-19 Patients on Day 0 in the ICU. Frontiers in Digital Health, 3. https://doi.org/10.3389/fdgth.2021.681608

Janezic, S., Mahnic, A., Kuhar, U., Kovač, J., Jenko Bizjan, B., Koritnik, T., Tesovnik, T., Šket, R., Krapež, U., Slavec, B., Malovrh, T., Battelino, T., Rupnik, M., & Zohar Cretnik, T. (2023). SARS-CoV-2 molecular epidemiology in Slovenia, January to September 2021. Eurosurveillance, 28(8). https://doi.org/10.2807/1560-7917.ES.2023.28.8.2200451

Krüger, L. J., Gaeddert, M., Tobian, F., Lainati, F., Gottschalk, C., Klein, J. A. F., Schnitzler, P., Kräusslich, H.-G., Nikolai, O., Lindner, A. K., Mockenhaupt, F. P., Seybold, J., Corman, V. M., Drosten, C., Pollock, N. R., Knorr, B., Welker, A., de Vos, M., Sacks, J. A., & Denkinger, C. M. (2021). The Abbott PanBio WHO emergency use listed, rapid, antigen-detecting point-of-care diagnostic test for SARS-CoV-2—Evaluation of the accuracy and ease-of-use. PLOS ONE, 16(5), e0247918. https://doi.org/10.1371/journal.pone.0247918

Landaas, E. T., Storm, M. L., Tollånes, M. C., Barlinn, R., Kran, A.-M. B., Bragstad, K., Christensen, A., & Andreassen, T. (2021). Diagnostic performance of a SARS-CoV-2 rapid antigen test in a large, Norwegian cohort. Journal of Clinical Virology, 137, 104789. https://doi.org/10.1016/j.jcv.2021.104789

Li, D., Zhang, Q., Tan, Y., Feng, X., Yue, Y., Bai, Y., Li, J., Li, J., Xu, Y., Chen, S., Xiao, S.-Y., Sun, M., Li, X., & Zhu, F. (2020). Prediction of COVID-19 Severity Using Chest Computed Tomography and Laboratory Measurements: Evaluation Using a Machine Learning Approach. JMIR Medical Informatics, 8(11), e21604. https://doi.org/10.2196/21604

Lorenzoni, G., Sella, N., Boscolo, A., Azzolina, D., Bartolotta, P., Pasin, L., Pettenuzzo, T., De Cassai, A., Baratto, F., Toffoletto, F., De Rosa, S., Fullin, G., Peta, M., Rosi, P., Polati, E., Zanella, A., Grasselli, G., Pesenti, A., Navalesi, P., … VENETO ICU Network. (2021). COVID-19 ICU mortality prediction: a machine learning approach using SuperLearner algorithm. Journal of Anesthesia, Analgesia and Critical Care, 1(1), 3. https://doi.org/10.1186/s44158-021-00002-x

Masiá, M., Fernández-González, M., Sánchez, M., Carvajal, M., García, J. A., Gonzalo-Jiménez, N., Ortiz de la Tabla, V., Agulló, V., Candela, I., Guijarro, J., Gutiérrez, J. A., de Gregorio, C., & Gutiérrez, F. (2021). Nasopharyngeal Panbio COVID-19 Antigen Performed at Point-of-Care Has a High Sensitivity in Symptomatic and Asymptomatic Patients With Higher Risk for Transmission and Older Age. Open Forum Infectious Diseases, 8(3). https://doi.org/10.1093/ofid/ofab059

Nagy, É., Golopencza, P., Barcs, I., & Ludwig, E. (2023). Comparison of COVID-19 Severity and Mortality Rates in the First Four Epidemic Waves in Hungary in a Single-Center Study with Special Regard to Critically Ill Patients in an Intensive Care Unit. Tropical Medicine and Infectious Disease, 8(3), 153. https://doi.org/10.3390/tropicalmed8030153

Najafi-Vosough, R., Bakhshaei, M. H., & Farzian, M. (2023). Predicting Mortality and ICUs Transfer in Hospitalized COVID-19 Patients Using Random Forest Model. Archives of Anesthesia and Critical Care. https://doi.org/10.18502/aacc.v9i6.14444

Ovcharenko, E., Kutikhin, A., Gruzdeva, O., Kuzmina, A., Slesareva, T., Brusina, E., Kudasheva, S., Bondarenko, T., Kuzmenko, S., Osyaev, N., Ivannikova, N., Vavin, G., Moses, V., Danilov, V., Komossky, E., & Klyshnikov, K. (2023). Cardiovascular and Renal Comorbidities Included into Neural Networks Predict the Outcome in COVID-19 Patients Admitted to an Intensive Care Unit: Three-Center, Cross-Validation, Age- and Sex-Matched Study. Journal of Cardiovascular Development and Disease, 10(2). https://doi.org/10.3390/jcdd10020039

Platten, M., Hoffmann, D., Grosser, R., Wisplinghoff, F., Wisplinghoff, H., Wiesmüller, G., Schildgen, O., & Schildgen, V. (2021). SARS-CoV-2, CT-Values, and Infectivity—Conclusions to Be Drawn from Side Observations. Viruses, 13(8), 1459. https://doi.org/10.3390/v13081459

Sakagianni, A., Koufopoulou, C., Verykios, V., Loupelis, E., Kalles, D., & Feretzakis, G. (2023). Prediction of COVID-19 Mortality in the Intensive Care Unit Using Machine Learning. https://doi.org/10.3233/SHTI230200

Shen, Q. (2023). Research of mortality risk prediction based on hospital admission data for COVID-19 patients. Mathematical Biosciences and Engineering : MBE, 20(3), 5333–5351. https://doi.org/10.3934/mbe.2023247

Shi, Y., Qin, Y., Zheng, Z., Wang, P., & Liu, J. (2023). Risk Factor Analysis and Multiple Predictive Machine Learning Models for Mortality in COVID-19: A Multicenter and Multi-Ethnic Cohort Study. The Journal of Emergency Medicine, 65(6), e584–e593. https://doi.org/10.1016/j.jemermed.2023.06.012

Treggiari, D., Piubelli, C., Caldrer, S., Mistretta, M., Ragusa, A., Orza, P., Pajola, B., Piccoli, D., Conti, A., Lorenzi, C., Serafini, V., Boni, M., & Perandin, F. (2022). SARS-CoV-2 rapid antigen test in comparison to RT-PCR targeting different genes: A real-life evaluation among unselected patients in a regional hospital of Italy. Journal of Medical Virology, 94(3), 1190–1195. https://doi.org/10.1002/jmv.27378

Viderman, D., Kotov, A., Popov, M., & Abdildin, Y. (2024). Machine and deep learning methods for clinical outcome prediction based on physiological data of COVID-19 patients: a scoping review. International Journal of Medical Informatics, 182, 105308. https://doi.org/10.1016/j.ijmedinf.2023.105308

Wagenhäuser, I., Knies, K., Rauschenberger, V., Eisenmann, M., McDonogh, M., Petri, N., Andres, O., Flemming, S., Gawlik, M., Papsdorf, M., Taurines, R., Böhm, H., Forster, J., Weismann, D., Weißbrich, B., Dölken, L., Liese, J., Kurzai, O., Vogel, U., & Krone, M. (2021). Clinical performance evaluation of SARS-CoV-2 rapid antigen testing in point of care usage in comparison to RT-qPCR. EBioMedicine, 69, 103455. https://doi.org/10.1016/j.ebiom.2021.103455

Zhan, C., Zheng, Y., Zhang, H., & Wen, Q. (2021). Random-Forest-Bagging Broad Learning System With Applications for COVID-19 Pandemic. IEEE Internet of Things Journal, 8(21), 15906–15918. https://doi.org/10.1109/JIOT.2021.3066575

Zhao, Y., Zhang, R., Zhong, Y., Wang, J., Weng, Z., Luo, H., & Chen, C. (2022). Statistical Analysis and Machine Learning Prediction of Disease Outcomes for COVID-19 and Pneumonia Patients. Frontiers in Cellular and Infection Microbiology, 12. https://doi.org/10.3389/fcimb.2022.838749